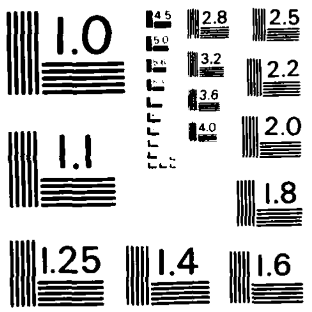
AD-A130 794 LOW ALTITUDE SIMULATOR TRAINING: A-10 AIRCRAFT(U) AIR FORCE HUMAN RESOURCES LAB BROOKS AFB TX B J PIERCE JUN 83 AFHRL-TP-82-43 1/1

UNCLASSIFIED F/G 1/2 NL

END
DATE
FILMED
8-83
DTIC


1.0
4.5
5.0
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.1
1.8
1.25
1.4
1.6

AFHRL-TP-82-43

(12)

AIR FORCE

HUMAN

RESOURCES

LABORATORY

LOW ALTITUDE SIMULATOR TRAINING: A-10 AIRCRAFT

By

Byron J. Pierce

OPERATIONS TRAINING DIVISION
Williams Air Force Base, Arizona 85224

June 1983

Final Technical Paper

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

MILTON E. WOOD, Technical Director
Operations Training Division

CARL D. ELIASON, Colonel, USAF
Chief, Operations Training Division

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFHRL-TP-82-43	2. GOVT ACCESSION NO. ADA130 794	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) LOW ALTITUDE SIMULATOR TRAINING: A-10 AIRCRAFT		5. TYPE OF REPORT & PERIOD COVERED Final 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Byron J. Pierce		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Training Division Air Force Human Resources Laboratory Williams Air Force Base, Arizona 85224		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62205F 11231115
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE June 1983 13. NUMBER OF PAGES 42
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS (of this report) Unclassified 15.a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of this abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

flying training
operational test, evaluation
personnel training, evaluation
pilot training
simulator training, evaluation
syllabus development

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Simulator scenarios were developed to train student pilots in A-10 aircraft low level navigation (LLN) and elementary basic attack maneuvers (BAMs). The primary objective of this effort was to evaluate the effectiveness of these scenarios using the A-10 configured Advanced Simulator for Pilot Training (ASPT) as the training medium. Training effectiveness was assessed using simulator automated performance measures and instructor pilot (IP) evaluations of student aircraft performances in a transfer of training paradigm. There were 42 subjects in the BAM phase of the experiment and 36 subjects in the LLN phase. Simulator measures showed enhanced performance as

DD Form 1473 1 Jan 73 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 20 (*Continued*)

a function of training fo: .hree of the five BAM tasks trained; LLN simulator performance data were not analyzed due to difficulty with the automated scoring procedures. Power values determined for aircraft performance measures and inferences based on the results of similar efforts made questionable the validity of using IP performance evaluations for test purposes as was done in this experiment. Discussion is directed toward future test programs requiring assessment of pilot performances in tactical aircraft.

AFHRL Technical Paper 82-43 June 1983

LOW ALTITUDE SIMULATOR TRAINING: A-10 AIRCRAFT

By

Byron J. Pierce

OPERATIONS TRAINING DIVISION
Williams Air Force Base, Arizona 85224

Reviewed and submitted for publication by

Thomas H. Gray
Chief, Operational Unit Training Branch

This publication is primarily a working paper.
It is published solely to document work performed.

A

PREFACE

This research and development effort was conducted by the Operations Training Division of the Air Force Human Resources Laboratory, Air Force Systems Command, in coordination with Headquarters, Tactical Air Command and was supported by the 355th Tactical Training Wing and the 333rd, 357th and 358th Tactical Fighter Training Squadrons at Davis-Monthan Air Force Base, Arizona. The effort was performed to satisfy requirements of Air Force Human Resources Laboratory Technical Planning Objective 3, Air Combat Tactics and Training. Work Unit 11231115, A-10 Combat Scenario Development and Evaluation, addressed a portion of this objective, namely, improved mission survival in combat training. Mr. James F. Smith was the project scientist.

The author extends his appreciation to the members of the 355th Tactical Training Wing and the 333rd, 357th, and 358th Tactical Fighter Training Squadrons for their interest and cooperation throughout this effort and in particular to Maj John Roulston, Maj James Cizek, and Capt Charles Stenner, without whose assistance the work could not have been completed. The author would also like to thank Mr. Richard Greatorex and Mr. Tien Fu Sun of the Operations Training Division for their advice and assistance in analyzing the data.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LOW ALTITUDE SIMULATOR TRAINING: A-10 AIRCRAFT

I. INTRODUCTION

This report is organized in the following manner. This section (Introduction) reviews background material, describes the experimental task selection process, provides a description of these tasks, and presents the objectives of the effort. Section II describes the experimental training apparatus and presents the general approach used for study. To facilitate operational scheduling and experimental design requirements, the effort was separated into two phases. Each phase is reported as a separate unit. Phase I described in Section III, focuses on experimental issues concerning training of elementary basic attack maneuvers (BAMs). Phase II deals exclusively with low level navigation (LLN) training issues and is described in Section IV. Subjects, experimental training procedures, dependent measures, data analysis procedures, and results are explained for each phase separately in their respective sections. Section V presents discussion of the results obtained from both phases. Finally, conclusions and recommendations are presented in Section VI.

Background

Tactical requirements to fly low altitude operations in any terrain and in any geographical area have been difficult to satisfy. The A-10 aircraft evaluation reports from Bentwaters AFB and Myrtle Beach AFB have shown that transition training B-course graduates typically have a minimum 6-month orientation period before they are considered ready to fly lead on low altitude operations. Data from actual combat situations suggest that orientation to novel low altitude environments may be a persistent problem for the first 10 sorties (Joss, 1978). The use of simulators for training low altitude operations represents an attractive possibility as a means for reducing these lengthy orientation periods.

Simulator research and development involving low altitude operations has focused primarily on visual cue requirements. Current computer-generated imagery (CGI) capabilities preclude simulation of highly detailed low altitude environments. Various coplanar and vertical pattern configurations have been evaluated in an attempt to optimize existing CGI capabilities. Reviews of this work can be found in Edwards, Pohlman, Buckland, and Stephens (1981) and Martin and Rinalducci (in press). The use of existing CGI capabilities in the development of operational scenarios for training low altitude procedures is an area of research not previously addressed. Whether simulator to aircraft transfer benefits can be derived from such scenarios is an issue of major interest to operational program managers. The current effort addressed this issue by examining the effectiveness of simulator scenarios developed to train elementary BAM and LLN tasks. The effort was conducted in conjunction with an ongoing A-10 Advanced Simulator for Pilot Training (ASPT) conversion and surface attack training program designed for initial entry students transitioning to the A-10 aircraft. The experiment was designed to have minimal impact on regular operations of the A-10 training program.

Experimental Tasks Selection

The selection of experimental tasks was based on two requirements. First, the tasks had to be normally trained in the aircraft at low to medium altitudes. The rationale supporting this requirement was the research need to address transfer effectiveness issues of tactical low altitude training in the simulator. Second, to properly evaluate the training effectiveness of the ASPT while minimizing scheduling conflicts with regular squadron operations, it was necessary to select tasks which were initially practiced in the aircraft shortly after scheduled ASPT sessions. Elementary BAM and LLN tasks fulfilled these requirements. Both tasks offered operations in the low to medium altitude regime and both tasks were initially trained in the A-10 aircraft shortly after the scheduled ASPT training sessions.

Experimental Tasks Description

Elementary BAM. Elementary BAMs are the building blocks used in practically all surface attack maneuvers. Elementary BAMs consist of 90°/180° low altitude turns, 30° and 20° dive, roll-in and recovery (DRR) tasks, and target tracking exercises. The purpose of elementary BAMs is to familiarize the student pilot with the use of the head-up display

(HUD) and provide experience in surface attack procedures. Specifically, 90°/180° turns are designed to train the pilot to execute a turn at a prescribed angle of attack (AOA), to maintain a prescribed altitude, to roll out approximately the prescribed number of degrees from the initial heading, and to maintain situation awareness. The 20° and 30° DRR tasks are designed to acquaint the pilot with techniques required for surface attack maneuvers. Additionally, the pilot is taught to establish a cross check of HUD data to include dive angle, airspeed, and altitude. Target tracking is required for gun employment. The objective of this exercise is to track the gun cross up to a target, hold it there momentarily, and be aware of parameters displayed in the HUD.

LLN. The objective of the LLN training mission is to develop the basic skills necessary to perform tactical navigation. Each LLN mission involves the review and execution of a predetermined navigation exercise to a target or practice area. The LLN route consists of a number of sections or legs. The student is provided topographic maps of each leg with headings, times, and distances clearly indicated. The LLN training mission is designed to instruct the student in the fundamentals of low level operations and dead reckoning navigation. Low-level fundamentals include route study, route timing, route entry, flight formation, radio procedures, navigation error analysis, and aircraft control techniques. Dead reckoning consists of flying from one point to another relying solely on time, speed, and direction. To successfully master dead-reckoning procedures, the pilot must be proficient at map reading, identifying visual checkpoint features, and determining the effect wind has on the aircraft ground track. LLN training given to the A-10 B-course pilot focuses on those aspects of the mission peculiar to the A-10 weapon system. After completion of this training, the pilot is taught more advanced operations involving low altitude tactical navigation and low altitude tactical formation.

Objectives

There were two primary objectives of this effort: (a) to evaluate the effectiveness of ASPT scenarios to train elementary BAMs, and (b) to evaluate the effectiveness of ASPT scenarios to train LLN. A transfer-of-training paradigm was used to address these objectives. Training effectiveness was operationally evaluated by examining trends in simulator performance data and by examining differences between experimental and control group performances on tasks performed in the aircraft.

A secondary objective to the experiment was to evaluate the sensitivity of the aircraft analysis variables in the detection of performance differences between experimental and control groups. IP evaluations were the only metrics available for assessing BAM and LLN student aircraft performances. Objective measures, such as bomb scores, were not obtainable by the very nature of these tasks. All aircraft performance analysis variables used in this effort were, therefore, based on IP evaluation data. The sensitivity of IP evaluations to between group performance variations has been shown to be much less than the sensitivity of more objective measures of performance for weapons delivery tasks (Gray & Fuller, 1977). Since this was seen as a potential problem, post hoc evaluations of statistical power were used to operationally evaluate the sensitivity of aircraft performance variables used in the analyses.

II. GENERAL APPROACH

The apparatus and methodology descriptions presented in this section are generic to both BAM and LLN phases. The methodology concerns that are specific to the individual phases are presented in their respective sections.

Training Apparatus

The A-10 ASPT was used in this experiment to train experimental group subjects BAM and LLN exercises. A complete description of the A-10 ASPT system can be found in Gray, Chun, Warner, and Eubanks (1981) and a summary of recent modifications is included in Brooks and Lyon (in press). The g-seat and g-suit force cueing systems on the ASPT were used in this effort; however, the ASPT platform motion system was not. The g-suit used in the experiment was the standard g-suit used in the A-10 aircraft. Inflation/deflation of the suit was computer controlled and simulated the suit pressures present during actual aircraft maneuvers. The visual system on the ASPT consisted of seven 36-inch monochromatic cathode ray tubes (CRTs) placed around the cockpit to allow the pilot +110° to −40° vertical cueing

and ±150° horizontal cueing. BAM and LLN visual environments were developed applying experience, research findings, and inferences from previous ASPT CGI development efforts.

Method

As part of the regular training syllabus, A-10 B-course student pilots received two training sessions in the ASPT. The first session, consisting of two 1.5-hour sorties, was used to train conversion tasks; surface attack maneuvers were trained in the second session, which consisted of three 1.5-hour sorties. Appendix A contains the A-10 ASPT syllabi used at the time of this experiment. The two training sessions were scheduled approximately 3 weeks apart. During this 3-week period, students received conversion phase training in the aircraft, along with scheduled academics. Students typically flew their first BAM aircraft sortie within 1 week following ASPT conversion training. It took approximately 10 days for all students within a class to complete the three required BAM aircraft sorties. Two aircraft sorties containing LLN exercises were scheduled within 1 week following ASPT surface attack training. The delay between training in the ASPT and the first of the two LLN aircraft sorties was typically 3 to 4 days. The separation of BAM and LLN operations into two phases allowed independent group assignment and data analysis procedures. This design had minimal impact on training operations and was acceptable to operational training personnel. Briefings for both BAM and LLN ASPT sorties were delivered by A-10 IPs. The same three A-10 IPs alternated briefing and instructing students on research sorties throughout the effort. On each sortie, students received conventional instruction and feedback from the IP. Except for the students themselves and the research IPs, student group assignments were unknown to squadron personnel.

III. PHASE I - BAM

Subjects

The subjects were 42 A-10 B-course student pilots assigned to the 355th Tactical Training Wing, Davis-Monthan AFB, served as subjects. All subjects were recent Undergraduate Pilot Training (UPT) graduates and had completed fighter lead-in training, but had not been qualified in any other operational aircraft. As part of the normal syllabus, each subject was a participant in the A-10 ASPT training program. For study purposes, these student pilots were randomly assigned to either an experimental or a control group; 22 were assigned to the experimental group, and the remaining 20 served as controls.

Experimental Training

Upon completion of ASPT conversion training, experimental group subjects received a briefing that covered the visual environment, task requirements, and procedures for the BAM experimental sortie. After this briefing, subjects flew the 1-hour BAM sortie on the ASPT with an A-10 IP providing instruction from the main console. The sortie was developed with the help of IPs from the A-10 operational training development section at Davis-Monthan AFB. The sortie began with a takeoff followed by a 5-minute area orientation period. Subjects then performed each task a set number of trials. A detailed syllabus of the BAM sortie is provided in Table 1. Due to the nature of the tasks being trained, the experimental sortie was flown in a continuous fashion; at no time during the training period were the simulator freeze or reset functions used. All control group subjects received a 1-hour ASPT "control" sortie following ASPT conversion training to equate the groups with respect to total time in the simulator. The control sortie was simply a 1-hour repeat of the last sortie flown during ASPT conversion training (Appendix A).

Table 1. **Detailed ASPT BAM Syllabus**

1a	Takeoff; Climb; Level off @ 6000 feet AGL (wx:2000/3)
1b	Area Orientation (5 minutes)
2a	90°/180° Turns (17.5 units AOA) - (3L/3R repetitions)
2b	90°/180° Turns (21.5 units AOA) - (3L/3R repetitions)
3	30° Dive, Roll-In and Recovery (7 repetitions)
4	20° Dive, Roll-In and Recovery (7 repetitions)
5	Target Tracking (5 repetitions)

The computer-generated visual environment used for ASPT BAM training during phase I was an area approximately 9 miles by 9 miles (Figure 1). In this area were placed several mountain ranges, ranging in height from 3500 feet to 5500 feet above ground level (AGL), with scattered 2000-foot-AGL pinnacles. These mountains surrounded a gaming area approximately 3 miles by 3 miles in size. Within this gaming area were 50-foot inverted tetrahedrons used to provide vertical development cues at low altitudes. These cues covered the entire gaming area with 1000-foot separation. All BAM tasks were practiced within this gaming area. For the target tracking task, one of several tank-shaped targets in the area was made twice its normal size and placed upon a 100-foot-square white background. This was done to guarantee ease in locating the correct target. None of the other targets located in the area had any special purpose in this sortie. A runway that was inserted into the environment at the beginning of the sortie was used for takeoffs. After the simulator indicated airborne at 2000 feet AGL, the visibility of the gaming area was reduced to 3 miles, and the runway was removed. At 6000 feet AGL, unlimited visibility was restored over the gaming area, and the exercise began.

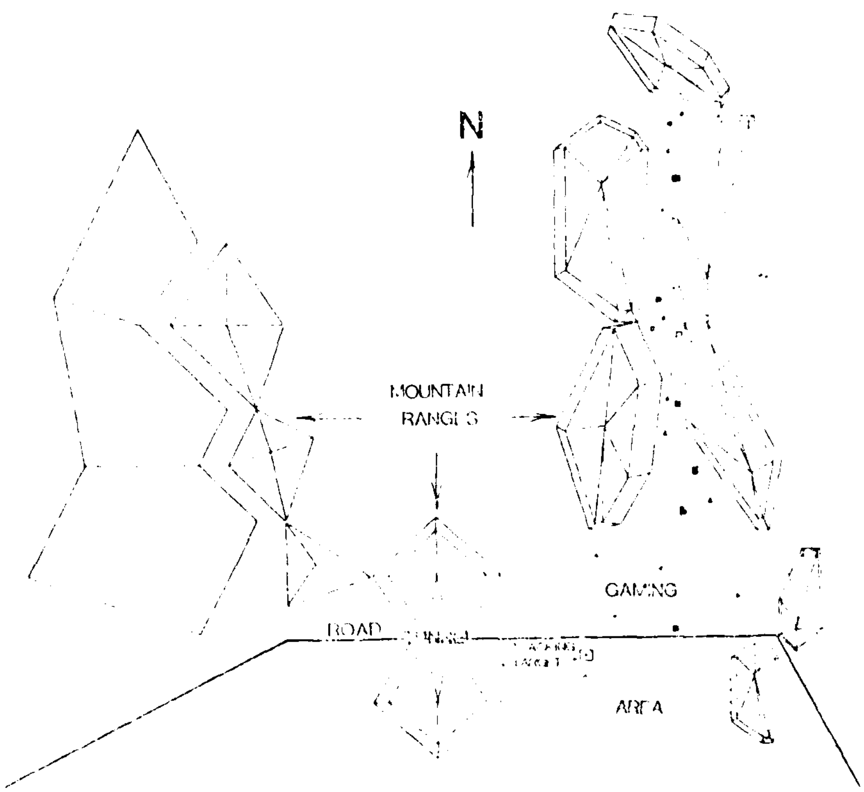


Figure 1. **Map representation of the ASPT BAM visual environment.**

Dependent Measures

ASPT Performance Measurement. BAM simulator performance data were collected on all experimental subjects using the simulator's automated performance measurement system (PMS). A description of the ASPT automated PMS can be found in Gray et al. (1981). The measures collected for each BAM task, along with the start/stop parameters used to define data collection periods, are listed in Table 2.

Table 2. **ASPT BAM Performance Measures**

Training Exercise	Scoring Logic: Start	Scoring Logic: Stop	Performance Measures
90°/180° Turns (AOA = 17.5 & 21.5) (Altitude = 500 feet)	ABS Bank GT 20°	ABS Bank LT 5°	RMS deviations from desired AOA RMS deviations from desired Altitude
30° and 20° DRR	30° event: ABS Bank LT 10°: Dive Angle GT 20° 20° event: ABS Bank LT 10°: Dive Angle GT 15°	Vertical Velocity GT 0 Vertical Velocity GT 0	RMS deviations from desired Dive Angle on Final RMS G-loading on Final Pickle Parameters (captured @ 2500 feet AGL for 30° and 1800 feet AGL for 20°): Bank Angle Dive Angle Knots Indicated Airspeed G loading
Target Tracking	Within 1800 feet of Target: Heading Between + 110° & −150°; ABS Bank LT 20°	Altitude LT 1400 feet AGL	Hit/Miss on target (captured between 1600 feet & 1400 feet AGL) RMS miss distance from target feet (between 1600 feet & 1400 feet AGL)

Aircraft Performance Measurement. Student performance in the aircraft was evaluated using data obtained from student mission gradesheets and BAM performance questionnaires developed for this effort (Appendix B, Table B-1). Flight IPs completed student gradesheets and performance questionnaires at the debriefing of each flight. Both measurement instruments used a 5-point rating scale. This rating scale was the standard 0 to 4 scale used within TAC for evaluating student performances: 0 ratings indicated lack of ability or knowledge; 1 ratings indicate safe but limited proficiency; 2 ratings indicated performance was essentially correct with errors identified and corrected; 3 ratings indicated performance was correct, efficient, skillful and without hesitation; and ratings of 4 reflected performance of an unusually high degree of ability. Rating criteria for individual tasks were specified in the A-10 Flight Objectives Pamphlet. All B-course conversion phase sorties were flown as two-ship flights composed of one student with one IP. Due to the operational environment and non-interference status of the effort, IP/student assignments could not be experimentally controlled. Assignments were usually based on IP availability and student progress.

The student mission gradesheets were used to assign an overall rating to each task performed on a flight. The A-10 B-course syllabus contained three sorties on which BAM tasks were practiced. Gradesheet ratings assigned to BAM tasks performed on these sorties were recorded for later analysis. Individual repetitions data were obtained using the BAM performance questionnaire. This questionnaire was designed to obtain measures similar to those obtained from the ASPT automated PMS. Using the questionnaire, IPs provided overall ratings on each BAM task repetition performed. In

addition, maximum altitude deviations from desired altitude were recalled for each 17.5 and 21.5 unit AOA turn performed. IPs also assigned AOA control ratings to each turn repetition. For 20° and 30° DRR tasks, IPs recalled deviations in airspeed and dive angle from desired parameters at pickle altitude for each repetition. Not all BAM tasks were performed on each of the three BAM aircraft sorties. Table 3 lists the tasks that were practiced during each flight.

Table 3. **Distribution of BAM Tasks**

Sortie	TASK			
	Turns	20° DRR	30° DRR	Target Tracking
CV-4	Yes	No	Yes	Yes
CV-5	Yes	Yes	No	No
CV-6	Yes	Yes	Yes	No

Squadron instructor pilots (IPs) were briefed on the purpose and requirements of Phase I of the program. Additionally, a letter outlining the data collection requirements was circulated among the participating squadrons prior to each data collection period. During these periods, a researcher was on site to administer questionnaires, collect completed data collection materials, and answer questions pertaining to the effort. Despite these efforts, aircraft performance data for two experimental subjects were removed from final analysis due to IP misunderstandings of data requirements. No problems were encountered with the aircraft performance data collected for the 20 control subjects and all of these subjects were used in the final analyses.

Data Analysis Procedures

Simulator performance data analyses focused on identifying and describing significant changes in performance as a function of training. Single factor with repeated measures (A x S) Analysis of Variance (ANOVA) procedures (Keppel, 1973) were used to analyze all intratrial RMS deviation data obtained for each task. The same ANOVA procedures were used to analyze transformed values of release altitude parameters for 20° and 30° (DRR) exercises. Post hoc comparisons on trial means were made using linear trend analyses procedures (Keppel, 1973) for variables whose ANOVAs were found to be significant at the $p < .05$ level.

Aircraft performance data analyses examined between group differences and identified those variables most sensitive to group performance variations. Aircraft performance data obtained from BAM performance questionnaire and student mission gradesheets were sorted into individual analysis variables for statistical examination. Selection of analysis variables for the various tasks was based on a requirement to achieve compatibility with the simulator performance analysis variables. In addition, it was desirable that these aircraft performance variables be similar to those outlined in test plans for the test and evaluation of future Tactical Air Command (TAC) weapon system trainers. Single factor ANOVA procedures (Keppel, 1973) were used to test between group differences of each analysis variable. Because the use of parametric procedures to analyze subjective evaluations might be debated, analysis variables were also tested using the non-parametric Mann-Whitney U-statistic (Siegel, 1956). Finally, post hoc determinations of statistical power (Keppel, 1973) were computed for each variable to assess the sensitivity of the measures to group performance variations.

Phase I Results

Results of the analyses of simulator and aircraft performance data are presented separately, with simulator data analyses reported first.

ASPT Trend Analyses

17.5 and 21.5 Unit AOA Turns. Two measures of training performance were derived from the data for both 17.5 and 21.5 unit (AOA) turns; thus, a total of four measures were extracted. For each event, the two measures were intratrial RMS AOA deviation and intratrial RMS altitude deviation.

Post hoc examination of the data revealed that the start/stop logic used to determine data collection "window" periods had biased the AOA deviation calculations toward higher RMS values. At roll-in and roll-out, when AOA is typically far below the desired value, AOA deviations were erroneously collected for an undetermined number of iterations and incorporated in intratrial RMS calculations (review scoring logic in Table 2). The effect this had on resulting AOA RMS deviation values was not consistent; therefore, these data were not analyzed. To the extent that the most critical periods to control altitude during a turn are at roll-in and roll-out, the window logic was deemed appropriate for examining altitude control. The analyses of the training performances for 17.5 and 21.5 unit AOA turns was therefore restricted to intratrial RMS altitude deviation data.

ANOVAs were computed on intratrial RMS altitude deviation data for the 17.5 and 21.5 unit AOA turn events.[1] Results for the 17.5 unit AOA event were significant ($F_{9,186} = 2.22$; $p < .05$). Linear trend comparisons on trial means also produced significant results ($F_{1,20.6} = 9.1$; $p < .001$); all other orthogonal trend components were not significant. These results, combined with an examination of trial means (Appendix B, Table B-1), show a linear decrease in altitude deviation as a function of trials.

The ANOVA computed for 21.5 AOA turns data was *not* significant. Examination of Appendix C, Table C-1 shows less than a 60-foot range among trial mean deviations. Additionally, trial means fall far below those initially obtained for the 17.5 AOA event even though the 21.5 AOA event is considered by pilots to be more demanding. The 21.5 AOA event was trained immediately following the 17.5 AOA event; therefore carry-over effects may be partly responsible for the inability to demonstrate improved performances as a function of trials.

30° DRR. Two sets of variables were extracted from data for the 30° DRR exercise. The first set consisted of RMS dive angle deviation and RMS g values computed for "on-final" parameters. On-final for a DRR task was defined as the period from roll-out/wings-level to when the aircraft achieves weapons release altitude. The second set of variables consisted of release altitude (pickle) values of dive angle, airspeed, g, and bank angle transformed to estimated bombing impact error distances. These transformations ignored pipper placement error because specific pipper placement was not a task requirement. In addition, the transformations were based on the trajectory of a low drag bomb.

ANOVAs were computed for both of the on-final parameter values from the first set of variables. Results for dive angle deviation values showed significant differences across trials ($F_{6,97} = 2.26$; $p < .05$). The linear component of the trend comparisons computed for trials was significant ($F_{1,16.1} = 9.11$; $p < .001$); all other orthogonal trend components were not significant. These results combined with an examination of trial means (Appendix C, Table C-1) show significant decreases in RMS dive angle error as a function of trials. The results of the ANOVA computed for on-final RMS g values were not significant. The RMS g value trial means (Appendix C, Table C-1) showed fairly good control of g over all trials. Thus, little improvement could be expected.

Pickle parameter values were transformed to equivalent miss distances, in feet, to enable arithmetical computations and comparisons among the four variables. In dive bomb exercises, errors can cancel each other, thus minimizing overall bombing error. For the DRR tasks, however, the objective was not to compensate but to minimize all parameter errors. Absolute values of the transformations were therefore used in deriving analysis variables. ANOVAs were computed for transformed values of each parameter with additional ANOVAs computed on two combinations of the transformed values. The first combination was a summation of dive angle, airspeed, and g error transformations to obtain a total longitudinal deviation value. The second combination was a summation of dive angle, airspeed, g, and bank angle error transformations to obtain a total system deviation value. ANOVA results for the transformed data were significant for dive angle deviations ($F_{6,105} = 3.08$; $p < .025$), airspeed deviations ($F_{6,105} = 2.93$; $p < .025$), total longitudinal

[1] Due to missing observations, trials 11 and 12 were deleted from the analyses of 17.5 and 21.5 AOA data. Remaining missing observations for this and following analyses were calculated using the procedure described in Kirk (1968, p. 146).

deviation values ($F_{6,105} = 4.83$; $p < .001$), and total system deviation values ($F_{6,105} = 5.6$; $p < .001$). Trend comparisons of the transformed values for each of these variables produced significant linear component results for trials; no other orthogonal trend component was found to be significant. The F values and probability levels of the linear component in the trend analyses are as follows: dive angle deviation ($F_{1,17.5} = 15.1$; $p < .01$); airspeed deviations ($F_{1,17.5} = 13.0$; $p < .01$); summed longitudinal deviation values ($F_{1,17.5} = 27.9$; $p < .001$); and total pickle deviation values ($F_{1,17.5} = 19.2$; $p < .001$). These results combined with an examination of the trial means data (Appendix C, Table C-1) show decreases in error transformation values as a function of trials for dive angle, airspeed, longitudinal, and total pickle parameter deviations.

The results of the ANOVAs for g and bank angle transformed values were not significant. Examination of the transformed g mean values in Appendix C, Table C-1 shows results similar to those obtained from analysis of the 30° DRR on-final g data. Transformed mean values for g were small and remained consistent over the training period. Examination of bank angle transformed mean values in Appendix C-1 also shows little improvement over training. During the course of this effort, there existed a slight left to right drift on the ASPT visual system which might have influenced the bank angle deviation results. The extent to which this drift problem confounded the data, however, was unknown.

20° DRR. The same two variable sets derived from the 30° DRR data were used to examine 20° DRR performances. Similar ANOVAs were computed for both sets of variables. The only significant finding was for transformed values of pickle bank angle data from variable set two ($F_{6,117} = 2.25$; $p < .05$). Trend analysis of these data revealed a non-significant linear component, but did show a significant quadratic component ($F_{1,19.5} = .09$; $p < .025$). The trial means of the transformed bank angle pickle values are presented in Appendix C, Table C-1. These results are not consistent with performance trends usually obtained for training tasks of this nature. It is possible that the visual drift problem, identified earlier, may have influenced training performances enough to produce these findings.

Two possible explanations exist for the overall inability to detect performance improvements during the course of training the 20° DRR task. First, since the task was trained immediately following the 30° DRR task, carry-over effects may be responsible. Although parameter deviations for 20° bombing events have greater effect on bombing impact error than they do for 30° bombing events, the 20° DRR task was highly similar and mechanically was no more difficult to perform than the 30° DRR task. Comparison between the on-final parameter trial means for the two tasks (Appendix C, Table C-1) shows initial deviation values for the 20° DRR task are below those initially obtained for the 30° DRR task; thus, this explanation seems plausible. An alternative explanation concerns the amount of training devoted to this task. Given the operational constraints imposed upon this effort, simulator training time had to be minimized. Perhaps additional training in the simulator would have resulted in task performance improvements that were more observable.

Target Tracking. Two analysis variables were derived to examine training performances for the target tracking exercise. Both concentrated on pipper placement error as the pilot passed through firing altitude. The first variable consisted of intratrial RMS aiming error in feet. Aiming error data used to compute these values were collected during the period that the pilot passed through a 1600-foot to 1400-foot AGL firing altitude window. Aiming error was defined as the distance between the target and actual pipper placement. The second variable was derived from a hit/miss register of pipper placement accuracy. A hit was obtained for a given trial whenever the pipper passed through the target during the firing altitude window period.

The RMS aiming error was analyzed using the same ANOVA procedure. Results of this test were not significant for trials. RMS aiming error trial means are presented in Appendix C, Table C-1. Performance variations across trials for the hit/miss measures were tested using the nonparametric Cochran Q statistic (Siegel, 1956).[2] Results of this test were significant ($Q = 13.55$; $df = 4$; $p < .01$). Percentages of actual hit versus miss data are presented by trial in Appendix C, Table C-2. These data show that the significance of the Cochran Q statistic was due to improved performances as a function of trials.

---

[2] In performing this test, missing observations were counted as misses. Trial 4 contained one such missing observation, and trial 5 contained three. This was deemed the most conservative approach to handling the mission observations problem.

One possible explanation for the discrepancy between these test results is that, once the subject placed the aiming pipper on the target at proper altitude, there was no real attempt to hold it there throughout the entire firing altitude period. In other words, the subjects, while still within the firing altitude window, may have pulled off target after achieving what they considered a hit. This would explain the increase in the number of hits, with no apparent improvement in RMS aiming errors as a function of trials.

Aircraft Performance Group Comparisons

17.5 & 21.5 Unit AOA Turns. Student mission gradesheets did not differentiate between 17.5 and 21.5 unit AOA turns; therefore, separate analyses of gradesheet ratings for the two events were not possible. BAM student performance questionnaires did differentiate between the two events and separate analysis variables for each event were analyzed (see Appendix C, Table C-3). Better experimental group performance means were obtained for 27 of the 31 analysis variables. Results indicated none of the differences obtained for the four variables showing poorer experimental group performances were significant ($p > .05$). Of the 27 analysis variables whose group means indicated better experimental group performances, two resulted in significant findings. These were (a) V_{16} − Mean AOA control ratings over CV-5 repetitions for 17.5 AOA turns ($F_{1,38} = 13.68$; $p < .001$; $U_{20,20} = 298$; $p < .01$ (two-tailed)); and (b) V_{21} − Mean AOA control ratings over CV-5 repetitions for 21.5 AOA turns ($F_{1,38} = 5.25$; $p < .02$; $U_{20,20} = 271$; $p < .05$ (two-tailed)). Power estimates for each of the analysis variables are presented in Appendix C, Table C-3. Aside from the two variables associated with significant ANOVA results and the four variables with negative group mean differences, power estimates ranged from 0 to .24.

30° DRR. Twelve analysis variables were derived from the 30° DRR data (see Appendix C, Table C-4). Group mean differences for 10 of the analysis variables showed better experimental group performances relative to the control group. ANOVA results of the two variables showing poorer experimental group performances were not significant ($p > .05$). Of the 10 other analysis variables whose group means indicated better experimental group performances, one resulted in significant findings. This analysis variable was V_2 − number of repetitions to proficiency ($F_{1,38} = 5.07$; $p < .03$; $U_{20,20} = 136$; $p < .07$ (two-tailed)). Power estimates for the analysis variables showing better experimental group performances, but which resulted in nonsignificant ANOVAs, ranged from 0 to .21.

20° DRR. Twelve analysis variables, similar to those derived from the 30° DRR data, were derived from the 20° DRR data (see Appendix C, Table C-5). Group mean differences for eight of the analysis variables showed better experimental group performances relative to the control group. One analysis variable showed no difference between groups (V_2 − number of repetitions to proficiency). ANOVA results for the three variables which showed better control group performance were not significant ($p > .05$). Test results of all other 20° DRR variables found one to be significant. This finding came from analysis of variable V_6 − mean of the overall repetitions ratings on CV-5 repetitions ($F_{1,38} = 4.29$; $p < .04$; $U_{20,20} = 269$; $p < .06$; (two-tailed)). Power estimates for those non-significant differences showing better experimental group performances ranged from 0 to .24.

Target Tracking. Five analysis variables were derived from the aircraft target tracking training data (see Appendix C, Table C-6). Group mean differences for all variables showed better experimental group performances relative to the control group. Although none of the variables produced significant ANOVA results at the $p = .05$ level, two approached significance. One of these two variables produced significant Mann-Whitney U results. These two variables were (a) V_4 − Mean of the Overall Repetition Ratings ($F_{1,38} = 3.13$; $p < .09$; $U_{20,20} = 260$; $p < .10$ (two-tailed)); and (b) V_5 − Gradeslip Ratings ($F_{1,38} = 3.3$; $p < .08$; $U_{20,20} = 266$; $p < .03$ (two-tailed)). Power estimates over all variables ranged from 0 to .32.

The evidence for training transfer of the target tracking task seems stronger than that obtained for the three previous exercises. With the number of analyses conducted for each of the experimental tasks, however, conclusions should be based on trends in the results rather than on individual findings of a few analyses which reach significant levels of probability. This issue is presented in greater detail in the discussion section.

IV. PHASE II - LLN

Subjects

The subjects for Phase II were 36 A-10 B-course student pilots assigned to the 355th Tactical Training Wing, Davis-Monthan AFB. All students had participated as either an experimental or a control group subject in Phase I of this effort. Each subject was a participant in the A-10 ASPT surface attack phase training program. Student pilots from Phase I control and experimental groups were assigned to one of two groups. Although squadron scheduling conflicts prohibited the use of strict counterbalancing procedures, each Phase II group was composed of an approximately equal number of subjects from each of the two Phase I groups. Sample size differences resulted due to the scheduling conflicts and ASPT time constraints. Fifteen subjects were assigned to the Phase II experimental group. The remaining 21 subjects served as controls. Time constraints also prohibited a control sortie being given to control group subjects as was done in Phase I.

Experimental Training

LLN simulator profiles were developed with the help of IPs from the A-10 operational training development section at Davis-Monthan AFB. Upon completion of ASPT surface attack training, experimental group subjects received a LLN briefing which covered the visual environment, task requirements, maps of the LLN routes, and procedures for the experimental sortie. After receiving this briefing, subjects flew two LLN routes on the ASPT with an A-10 IP providing instruction from the main console.

Limitations in the image generation capabilities of the ASPT prohibited development of a highly detailed visual environment. In addition, development of a visual environment which simulated the actual routes used for training in the aircraft was prohibited due to time and resource constraints. The LLN simulator profiles therefore emphasized procedural training of tasks relevant to LLN operations. The visual environment used in these profiles was selected on the basis of area size and visual cue requirements of the LLN exercise. The visual environment selected was an accurately scaled representation of the Fulda, West Germany area. The main area of use was approximately 150 miles × 150 miles. Actual maps of the Fulda, West Germany, area were used to place mountains, roads, towns, and other major landmarks throughout the simulated environment. Elevations deviated from those indicated on the real maps of the area, but distances between landmarks were to scale. The ground level was 1000 feet above mean sea level (MSL) throughout the simulated environment except for mountain terrain areas.

A highly detailed runway was developed at Fulda and used as home base. Two route entry points were positioned close to this home base. From each entry point, low level corridors were constructed extending into the target area. One of these corridors consisted of three legs; the other corridor consisted of four legs. Routes 1 and 2 were mirror images of each other; i.e., ingress legs for route 1 served as egress legs for route 2 and vice versa. The average length of each leg was about 25 miles. Numerous trees, buildings, grain silos, and towns were used for low altitude visual cues within the low level corridors. The density of these cues was satisfactory for sustained flight at or above 300 feet AGL. The visual cues were scattered within 2 miles either side of the centerline of each leg, giving the corridors a width of approximately 4 miles. Towns, factories, road intersections, towers, and mountains were used as checkpoints. Representations of tanks were used as targets in the target area. The total flying distance from home base to the target area was about 110 miles on either corridor.

Students navigated the routes using topographic maps detailing the CGI environment displayed on the ASPT visual system. The maps were color copier reproductions of photographic maps detailing the low level corridors. The photographic maps were developed from an original artwork mapping all landmarks generated within the environment. Care was taken to obtain correct coloring and scaling of the final map reproductions. The maps were bound together to form separate low level map booklets of each route with parameter information displayed in a format identical to what is used in aircraft LLN operations.

Each exercise began with a takeoff from the Fulda airport. The student was vectored to the LLN route entry point that marked the beginning of the navigation exercise. The student flew the low level route to the target area, where strafe tasks were performed using tanks as targets. Upon completion of the strafe tasks, the student flew to the egress entry point and followed a low level route back to Fulda. The only real differences between the two routes were the leg headings flown to navigate the routes and the strafe approaches performed in the target area (on route 1, subjects performed conventional approaches to strafe; on route 2, subjects performed pop-up approaches). Detailed syllabi of the two sorties are provided in Table 4.

Table 4. **Detailed LLN ASPT Syllabus**

Route 1

1. Takeoff; Climb; Level off @ 6000 feet AGL (WX: 27,000 feet; Winds; 25 knots gusting to 30 knots @ 360°)
2. Vector to Checkpoint A
3. LLN Ingress: (Altitude – 500 feet AGL; Airspeed – 300 knots)

   a. Checkpoint A: Intersection of Three Roads (Heading – 091°)
   b. Checkpoint B: Railroad Depot (Heading – 151°)
   c. Checkpoint C: Town W (Heading – 039)
   d. Checkpoint D: Factory; Target Identification Point

4. Long Range Strafe (Conventional – 3 repetitions)
5. Two Target Strafe (Conventional – 3 repetitions)
6. LLN Egress: (Altitude – 500 feet AGL; Airspeed – 300 knots)

   a. Checkpoint D: Factory (Heading – 224°)
   b. Checkpoint E: Town W (Heading – 274°)
   c. Checkpoint F: Town X (Heading – 222°)
   d. Checkpoint G: Town Y (Heading – 311°)
   e. Checkpoint H: Town Z

7. Return to Fulda AAF

Route 2

1. Takeoff; Climb; Level off @ 6000 feet AGL (WX: 27,000 feet; Winds; 25 knots gusting to 30 knots @ 360°)
2. Vector to Checkpoint A
3. LLN Ingress: (Altitude – 500 feet AGL; Airspeed – 300 knots)

   a. Checkpoint A: Town Z (Heading – 117°)
   b. Checkpoint B: Town Y (Heading – 039°)
   c. Checkpoint C: Town X (Heading – 094°)
   d. Checkpoint D: Town W (Heading – 042°)
   e. Checkpoint E: Factory: Target Identification Point

4. Long Range Strafe (Pop-up – 3 repetitions)
5. Two Target Strafe (Pop-up – 3 repetitions)
6. LLN Egress: (Altitude – 500 feet AGL; Airspeed – 300 knots)

   a. Checkpoint E: Factory (Heading – 224°)
   b. Checkpoint F: Town W (Heading – 332°)
   c. Checkpoint G: Railroad Depot (Heading – 269°)
   d. Checkpoint H: Intersection of Three Roads

7. Return to Fulda AAF

For both routes, winds were set at 20 knots, gusting to 30 knots, with delays of 15 to 30 seconds between gusts. A 5-second buildup time to maximum gust was programmed, and the duration of the gusts varied from 15 to 30 seconds. Wind direction varied ± 10 degrees from true north. Direction variations were made, on average, every 20 seconds. Turbulence was simulated at a moderate level and visibility was set at 4 1/2 miles.

Dependent Measures

ASPT Performance Measurement. Difficulties with the software logic used to capture ASPT performance measures were encountered throughout Phase II. This problem made questionable the validity of all ASPT LLN performance data, and analyses of these data are not reported. The measurement difficulties did not affect the quality of the actual LLN simulation flown by subjects. All experimental group subjects flew the entire LLN simulator profile. The training transfer issue for this exercise was addressed solely by between group aircraft performance analyses.

Aircraft Performance Measurement. Student data on LLN aircraft performances were obtained using procedures similar to those used for collecting BAM aircraft performance data in Phase I of this effort. The procedure incorporated both LLN performance questionnaires developed for this effort and student gradesheets. Both measurement instruments were completed by flight IPs at the debriefing of each flight. All LLN missions were flown with the student flying lead and the IP flying chase. As in Phase I, IP/student assignments were usually based on IP availability and student progress. Each student was scheduled to fly two aircraft sorties containing LLN exercises (BAM-1 and SA-1). Not all subjects, however, performed the scheduled number of LLN missions. Three subjects from the experimental group and four subjects from the control group had missing data for one LLN exercise due to weather restrictions on the day of flight. Complete sets of LLN aircraft performance data were collected for the remaining 29 subjects. All data obtained were used in the final analyses.

The student mission gradesheets used for LLN missions differed from the conversion phase gradesheets only on the tasks listed for rating. The procedures and the rating scale used were identical. IPs also completed the LLN student performance questionnaire presented in Appendix B, Table B-2 for each exercise performed. Using this questionnaire, IPs provided the following data for each leg of the exercise: (a) number of times lost or disoriented, (b) time deviation on checkpoint, (c) average altitude deviation, (d) average airspeed deviation, and (e) rating on course control. Based on IP interviews and review of mission objectives, these measures were determined to reflect the most important performance aspects of the LLN mission. The same procedures used to brief IPs on experiment requirements in Phase I of this experiment were also used in the LLN phase. Likewise, a researcher was on site during the data collection period to administer questionnaires, to collect completed measurement materials, and to answer questions pertaining to the effort.

Data Analysis Procedures

As previously noted, due to questionable validity, analyses of the LLN ASPT performance data are not reported. Aircraft performance data were analyzed in the following manner. The seven analysis variables presented in Appendixes D, Tables D-1 and D-2 were derived for each of the two LLN sorties. Selection of analysis variables was based on a desire to derive variables most reflective of mission performance and to achieve compatibility with simulator performance analysis variables planned for in the original design. For five of these variables, means were computed over data obtained for each route leg in order to make compatible data from different routes with varying number of legs. Single factor ANOVA and unequal sample size procedures (Keppel, 1973) were used to test between group differences for each variable. Post hoc determination of statistical power (Keppel, 1973) was also computed for each variable to assess the sensitivity of the data to group performance variations.

Phase II Results

Group mean differences for three of the seven BAM-1 sortie analysis variables showed better experimental group performances relative to the control group (see Appendix D, Table D-1). None of these differences were significant (p

> .05). Four variables showed better control group performances. Statistical analysis of these data found group differences for mean airspeed deviations (V_5) to be significant ($F_{1,31} = 10.25$; $p < .01$); the results of the analyses for the remaining three variables were not significant ($p > .05$). Group means for all SA-1 sortie variables indicated better control group performances (see Appendix D, Table D-2). Results of the statistical analyses of these data found none of these differences to be statistically significant.

V. DISCUSSION

Significant task improvements over BAM simulator periods were demonstrated for three of the five tasks trained. Analyses of the aircraft performance data for these three tasks resulted in significant positive findings for one of the 17 analyses computed for the 17.5 AOA turns task, one of the 12 analyses computed for the 30° DRR task, and two of the five analyses computed for the target tracking task. Similar results were obtained from aircraft performance analyses of the remaining BAM tasks: one of 17 analyses for 21.5 AOA turns data was significant, and one of 12 analyses for 20° DRR data was significant. Although few of the aircraft data analyses resulted in significant findings, group means obtained for these data showed an overall trend indicative of better performance by the simulator trained group. Simulator task performances could not be examined over the course of LLN training due to apparatus failure. Analyses of the LLN aircraft performance variables were not indicative of positive simulator to aircraft transfer of training.

Two factors which may have influenced the results of this effort are presented for consideration. The first factor concerns the adequacy of the training procedures. There are three components to this factor. The first component concerns the fact that the simulator scenarios for all tasks were, by necessity, restricted. The scope of the effort did not allow the determination of performance ceiling effects for task training in the simulator. Even for BAM tasks whose simulator performance data showed significant improvement as a function of training, the effect of additional training in the simulator on either simulator or aircraft task performances is unknown. The second component concerns the period between training in the simulator and aircraft testing. This period was not immediate: approximately 1 week elapsed between BAM training in the ASPT and the first BAM aircraft sortie; 3 to 5 days typically elapsed between the ASPT LLN sorties and the first LLN aircraft sortie. The third component is concerned primarily with LLN training operations. The ASPT's LLN visual environment was rather impoverished due to limitations of the ASPT CGI capability. None of the experimental pilots became lost during the course of performing LLN simulator scenarios. Additionally, the ASPT LLN visual environment was a representation of an area dissimilar to the real environment used for training in the aircraft. As a result, the ASPT LLN scenarios provided what might be termed generic procedural training; by no means were they intended to provide visual familarization of the test environment used for training in the aircraft. Although these issues may have negatively influenced the results of this effort, lack of complete control over training operations is commonplace in operational settings. The effort was designed to provide results based on anticipated usage of such training programs by the operational command. As such, the effort succeeded in accomplishing this objective while still maintaining the control over experimental procedures required by the methodology and analysis procedures used.

The second factor to be considered involves the sensitivity of the dependent variables used to detect aircraft task performance variations. Previous ASPT experiments involving weapons delivery training have demonstrated significant transfer to the aircraft when the dependent variables are based on objective measures of aircraft performance. Gray et al. (1981) used bombing error and percent strafe hits to demonstrate that, for five of six weapons delivery tasks tested, A-10 B-course pilots trained on the ASPT performed significantly better on initial A-10 surface attack aircraft sorties than did control counterparts. Similar results were obtained by Gray and Fuller (1977). In this experiment, fighter lead-in student pilots were trained on weapons delivery tasks in the ASPT and tested against control pilots in the F-5B aircraft. The aircraft performance variables used in this effort were derived from bombing error and IP evaluations of performances. Bombing error variables showed significantly better experimental group performances for all weapons delivery tasks tested; however, none of the group differences obtained for the IP evaluation variables were significant. As target tracking, 20° DRR and 30° DRR tasks are all similar in nature to their corresponding weapons delivery tasks, these previous efforts suggest that stronger evidence supporting transfer of training may have resulted if more objective aircraft performances metrics were available. The measurement sensitivity issue cannot be conclusively resolved here; however, power values determined for aircraft performance data analyses and inferences made on the basis of results from previous efforts make questionable the use of conventional IP evaluation procedures for test and evaluation purposes.

IP evaluations of student performances have long been used by TAC in determining individual pilot proficiency and, in general, the Command's overall combat readiness. The standardization/evaluation (STAN/EVAL) program serves the function of providing commanders with meaningful indicators that reflect aircrew training and capability to perform the unit mission. To achieve these ends, the STAN/EVAL program attempts to control the evaluation process and ensure standardization among the training units. In addition, criteria for IP assignment of student performance ratings are specified by criterion referenced objectives (CROs). One inherent problem is that, even by specifying the rating criteria, this evaluation process has no provision for verifying the subjective values produced. Additionally, the effect that confounding factors have on the final evaluations is difficult to ascertain. A student's previous performance history, attitude, and intentions may all significantly influence the subjective performance values assigned. As part of the STAN/EVAL program, students are measured against expected competence levels annotated in course syllabi. To this extent, the STAN/EVAL program itself may influence IP evaluations with factors unrelated to student performance.

On-going and planned procurements of new weapon system operational flight trainers (OFTs) for TAC pilot training have highlighted the need for evaluating these trainers in terms of their contribution to flying training programs. Test and evaluations are planned for these trainers using methodologies and subjective performance metrics similar to those used in the current effort. Examination of the power estimates presented in the appendixes provide insight to the kind of results to be expected unless another approach is taken. Power estimates can be interpreted in the following manner. Given the sample size and variability of the data between and within groups for a given measure, the probability of obtaining significant differences between groups is equal to the obtained power value. The power estimates presented in the appendixes were derived from mean squared data from the computed ANOVAs and used the F distribution in determining the power values. As a result, the power estimates are conservative, by nature of the F distribution. Even when more liberal methods are used, however, the resulting power estimates are quite moderate. Power increases as sample sizes increase, between group variability increases or the within group variability decreases. Large sample sizes are difficult to obtain for operational testing purposes. A desirable alternative to increasing sample sizes for purposes of increasing power would be to use measurement techniques that minimize sources of unwanted variability. The feasibility of adapting alternative measurement systems for aircraft performance assessments needs to be evaluated and, if suitable, incorporated into future simulator test and evaluation programs.

VI. CONCLUSIONS AND RECOMMENDATIONS

Two factors have been considered that may have negatively affected the outcome of this effort. First, training in the simulator was too restricted to ensure adequate transfer effects. Second, conventional IP evaluation procedures were not sufficiently sensitive for detecting group performance differences in the aircraft. Whether the obtained results were due to one, neither, or a combination of these proposals is unknown. The implications of these results form the basis of the following recommendations for both current and future training evaluation programs.

1. Ensure the optimal amount of benefit is derived from the similator if transfer of training is an issue. This can be accomplished by pretesting pilots in the simulator for performance ceiling effects, then programming enough simulator time to train pilots to optimal proficiency during the test period.

2. Evaluate the metrics used to assess task performances in terms of suitability and possible effects of unwanted sources of variance. When available, use objective performance measures. If subjective performance evaluations are necessary, rater instruction sessions and pretest procedures should be used to help ensure interrater reliability. Minimizing the number of raters used in the test and counterbalancing raters with ratees are also recommended.

REFERENCES

Brooks, R.B., & Lyon, D.R. *Force cue requirements for A-10 simulator weapons delivery training.* Technical Report. Williams AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory, in press.

Edwards, B.J., Pohlman, G.H., Buckland, G.H., & Stephens, C.W. Training low level terrain flight in a simulator. *Proceedings of the 3rd Interservice/Industry Training Equipment Conference and Exhibition*, Orlando, Florida, November 1981.

Gray, T.H., & Fuller, R.R. *Effects of simulator training and platform motion on air-to-surface weapons delivery training.* AFHRL-TR-77-29, AD-A043 649. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, July 1977.

Gray, T.H., Chun, E.K., Warner, H.D., & Eubanks, J.L. *Advanced flight simulator: Utilization in A-10 conversion and air-to-surface attack training.* AFHRL-TR-80-20, AD-A094 608. Williams AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory, January 1981.

Joss, J. Red flag: Realism on the range. *Air Force Magazine,* August 1978, **61**(8), 40–44.

Keppel, G. *Design and analysis: A researcher's handbook.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1973.

Kirk, R.E. *Experimental design: Procedures for the behavioral sciences.* Monterey, CA: Brooks/Cole, 1968.

Martin, E.L., & Rinalducci, E.J. *Effects of vertical cues on simulated low altitude flight: Study II.* Technical Report, Williams AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory, in press.

Siegel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw Hill, 1956.

APPENDIX A: A-10 ASPT SYLLABUS

Table A-1. **A-10 ASPT Detailed Conversion Phase Syllabus**

AS-1 — 1:1 Ratio
Simulator — 1.5 Hours

1. Takeoff; Climb; Level Off
2. Stall Recovery Procedures (Familiarization Only)

    a. Vertical Confidence Maneuver
    b. Power on Stalls (Straight Ahead and Turning)
    c. Traffic Pattern Stall Series
    d. Slow Flight

3. Lazy Eight; Aileron Rolls (Speed Brakes Closed and Speed Brakes Open – Familiarization Only)
4. Loop, Cuban Eight; Split S
5. Descent, Straight-In Approach; Go Around
6. Re-Entry; Normal Overhead Pattern and Approach; Go Around
7. Closed Pattern; Normal Approach; Touch and Go (Familiarization Only)
8. Closed Pattern; Normal Approach; Touch and Go (Familiarization Only)
9. Closed Pattern; Normal Approach; Touch and Go (Familiarization Only)
10. Closed Pattern; Normal Approach; Touch and Go (Familiarization Only)
11. Closed Pattern; Normal Approach; Touch and Go (Familiarization Only)
12. Re-Entry; Normal Overhead Pattern and Approach; Full Stop Landing

AS-2 — 1:1 Ratio
Simulator — 1.5 Hours

1. Takeoff; Climb; Level Off
2. Steep Turns
3. Re-Entry; Straight-In Approach (GCA); Go Around
4. Closed Pattern; Normal Approach; Touch and Go (Familiarization Only)
5. Closed Pattern; No Flap Approach; Touch and Go (Familiarization Only)
6. Closed Pattern; No Flap Approach; Touch and Go (Familiarization Only)
7. Straight-In Approach (Simulated Single Engine); Simulated Single Engine Go Around
8. Straight-In Approach (Simulated Single Engine); Touch and Go (Familiarization Only)
9. Closed Pattern; Normal Approach; Full Stop Landing
10. Takeoff (30); Departure
11. 30 Re-Entry

Table A-2. **A-10 ASPT Detailed Surface Attack Syllabus**

AS-3	1:1 Ratio
Simulator	1.5 Hours

1. Takeoff; Climb; Level Off
2. Range Orientation and Dive Bomb Demonstration (Familiarization Only)
3. 30° Dive Bomb (HOT) – 3 repetitions
4. 45° High Angle Dive Bomb (HOT) – 3 repetitions
5. Low Angle Low Drag (HOT) – 3 repetitions
6. Low Angle Bomb (HOT) – 3 repetitions
7. Low Angle Strafe (HOT) – 5 repetitions
8. Range Departure
9. Hung Ordnance Pattern (Runway 12); Full Stop Landing

AS-4	1:1Ratio
Simulator	1.5 Hours

1. Takeoff; Climb, Level Off
2. Range Entry
3. 30° Dive Bomb (HOT) – 3 repetitions
4. 45° High Angle Dive Bomb (HOT) – 3 repetitions
5. Low Angle Low Drag (HOT) – 3 repetitions
6. Low Angle Bomb (HOT) – 3 repetitions
7. Low Angle Strafe (HOT) – 5 repetitions
8. High Angle Strafe (HOT) – 5 repetitions
9. Range Departure
10. Hung Ordnance Pattern (Runway 30); Full Stop Landing

AS-5	1:1 Ratio
Simulator	1.5 Hours

1. Takeoff; Climb; Level Off
2. Range Entry
3. Low Angle Bomb Curvilinear (HOT) – 2 repetitions
4. Low Angle Bomb Pop-Up Demonstration – 1 repetition (Familiarization Only)
5. Low Angle Bomb Pop-Up (DRY) – 1 repetition
6. Low Angle Bomb Pop-Up (HOT) – 6 repetitions
7. Low Angle Strafe Pop-Up Demonstration – 1 repetition (Familiarization Only)
8. Low Angle Strafe Pop-Up (DRY) – 1 repetition
9. Low Angle Strafe Pop-Up (HOT) – 5 repetitions
10. Range Departure
11. Normal Overhead Pattern and Approach; Full Stop Landing

APPENDIX B: AIRCRAFT STUDENT PERFORMANCE QUESTIONNAIRES

Table B-1. **Elementary BAM Evaluation Form**
(to be completed for sorties CV4 through CV6)

Squadron____________________ Date____________________ Mission____________________

Pilot Name____________________________ IP Name____________________________

Event:	Results by Trial				
	1	2	3	4	5
1. 30° Dive, Roll and Recovery					
A/S Deviation @ 2500 ft (+ or −)	____	____	____	____	____
Dive Angle Deviation @ 2500 ft (+ or −)	____	____	____	____	____
Overall Rating (0–4)	____	____	____	____	____
2. 20° Dive, Roll and Recovery					
A/S Deviation @ 1800 ft (+ or −)	____	____	____	____	____
Dive Angle Deviation @ 1800 ft (+ or −)	____	____	____	____	____
Overall Rating (0–4)	____	____	____	____	____
3. Target Tracking					
Overall Rating	____	____	____	____	____
4. 17.5 AOA Turns (90/180)					
Max Altitude Deviation (+ or −)	____	____	____	____	____
AOA Control (0–4 Rating)	____	____	____	____	____
Overall Rating (0–4 Rating)	____	____	____	____	____
5. 21.5 AOA Turns (90/180)					
Max Altitude Deviation (+ or −)	____	____	____	____	____
AOA Control (0–4 Rating)	____	____	____	____	____
Overall Rating (0–4 Rating)	____	____	____	____	____

Table B-2. **A-10 Low Level Navigation Evaluation Form**
(to be completed at debriefing for sorties BAM-1 and SA-1)

Squadron ____________________ Date ____________________ Mission # ____________________

Pilot Name ______________________________ IP Name ______________________________

Route # ____________________ Entry Point ____________________ Exit Point ____________________

	Leg: 1	2	3	4	5
1. Number of Times Lost (disoriented)	____	____	____	____	____
2. Time Deviation on Checkpoint (Secs – or –)	____	____	____	____	____
3. Average Altitude Deviation (– or –)	____	____	____	____	____
4. Maximum Airspeed Deviation (– or –)	____	____	____	____	____
5. Course Control (0 – 4 Rating)	____	____	____	____	____

APPENDIX C: PHASE I TASK VARIABLES, GROUP MEANS AND AIRCRAFT DATA F VALUES WITH STATISTICAL POWER ESTIMATES

Table C-1. **ASPT BAM Variable Trial Means**

Maneuver	Variable	Trial Means 1	2	3	4	5	6	7	8	9	10
1. 17.5 Unit AOA Turns	RMS Altitude Deviation	228.2	214.9	195.8	183.0	180.3	205.0	113.2	152.0	143.7	108.7
2. 21.5 Unit AOA Turns	RMS Altitude Deviation	137.7	150.0	165.4	144.9	131.3	136.0	111.6	107.9	131.4	140.1
3. 30° DRR	**On Final Parameters:**										
	RMS Dive Angle Deviation	3.76	4.02	3.20	2.80	3.00	2.34	2.20			
	RMS G Values (desired: .88)	1.12	1.06	1.00	1.05	1.06	1.05	1.01			
	Pickle Parameters:										
	Dive Angle Deviation	108.9	94.5	64.4	69.7	72.9	48.0	41.4			
	Airspeed Deviation	81.0	85.7	62.4	57.3	58.4	61.4	59.4			
	G-Value Deviation	49.4	46.1	49.0	55.7	43.3	38.6	35.2			
	Longitudinal Total Error	239.3	226.3	175.8	182.7	174.6	148.0	136.0			
	Bank Angle Deviation	63.0	73.0	53.4	57.4	44.2	36.1	46.7			
	Total System Error	302.3	299.3	229.2	240.1	218.8	184.1	182.7			
4. 20° DRR	**On Final Parameters:**										
	RMS Dive Angle Deviation	2.81	2.47	2.23	2.88	2.76	2.12	2.44			
	RMS G Values (desired: .94)	1.13	1.12	1.24	1.13	1.14	1.12	1.06			
	Pickle Parameters:										
	Dive Angle Deviation	81.4	72.6	43.0	79.9	84.1	57.2	76.5			
	Airspeed Deviation	90.0	93.4	68.6	67.9	55.9	79.7	79.5			
	G Value Deviation	42.0	45.5	66.8	76.8	79.3	41.1	64.8			
	Longitudinal Total Error	213.4	211.5	178.4	224.6	219.3	178.0	220.8			
	Bank Angle Deviation	44.7	44.2	74.0	79.1	54.4	49.5	36.7			
	Total System Error	258.1	255.7	254.4	303.7	273.7	227.5	257.5			
5. Target Tracking	RMS Aiming Error	135.1	153.4	144.8	215.0	131.6					

Table C-2. **Percentage of Hits by Trial for Target Tracking Task**

	Trial				
	1	2	3	4	5
Hits, Percent	.45	.55	.64	.76	.84
Number of Observations[a]	22	22	22	21	19

[a]Fluctuations in numbers of observations are due to performance measurement system failures.

Table C-3. **Variables, Group Means and Statistical Power Estimates (alpha = .05) for Aircraft Turns Data**

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,38)	Power Estimate
V_1 - Highest Repetition Rating obtained in A/C (17.5 Turns)	3.1	3.05	.20	*
V_2 - Highest Repetition Rating obtained in A/C (21.5 Turns)	3.25	3.05	2.27	.21
V_3 - Number of Repetitions to Proficiency (17.5 Turns)	1.45	1.1	3.65	**
V_4 - Number of Repetitions to Proficiency (21.5 Turns)	1.35	1.45	.15	*
V_5 - 1st Repetition Overall Rating on CV-4 (17.5 Turns)	2.0	2.06	.23	**
V_6 - Mean of Altitude Deviations on CV-4 Repetitions (17.5 Turns)	108.04	123.14	.50	*
V_7 - Mean of AOA Control Ratings on CV-4 Repetitions (17.5 Turns)	2.49	2.41	.25	*
V_8 - Mean of Overall Ratings on CV-4 Repetitions (17.5 Turns)	2.41	2.36	.10	*
V_9 - CV-4 Gradeslip Ratings +	2.25	2.11	.64	*
V_{10} - 1st Repetition Overall Rating on CV-4 (21.5 Turns)	2.1	2.11	.01	**
V_{11} - Mean of Altitude Deviations on CV-4 Repetitions (21.5 Turns)	113.95	164.77	2.73	.24
V_{12} - Mean of AOA Control Ratings on CV-4 (21.5 Turns)	2.63	2.53	.39	*
V_{13} - Mean of Overall Ratings on CV-4 Repetitions (21.5 Turns)	2.47	2.35	.43	*
V_{14} - 1st Repetition Overall Rating on CV-5 (17.5 Turns)	2.28	2.11	.90	*
V_{15} - Mean of Altitude Deviations on CV-5 (17.5 Turns)	97.02	114.35	.35	*

Table C-3. (*Continued*)

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,38)	Power Estimate
V_{16} - Mean of AOA Control Ratings on CV-5 (17.5 Turns)	2.96	2.46	13.68***	.93
V_{17} - Mean of Overall Repetition Ratings on CV-5 (17.5 Turns)	2.59	2.44	1.63	.12
V_{18} - CV-5 Gradeslip Ratings +	2.6	2.4	1.58	.11
V_{19} - 1st Repetition Overall Rating on CV-5 (21.5 Turns)	2.5	2.26	1.63	.12
V_{20} - Mean of Altitude Deviations on CV-5 (21.5 Turns)	82.98	115.2	1.44	.09
V_{21} - Mean of AOA Control Ratings on CV-5 (21.5 Turns)	2.9	2.52	5.25***	.52
V_{22} - Mean of Overall Repetition Ratings on CV-5 (21.5 Turns)	2.64	2.59	.21	*
V_{23} - 1st Repetition Overall Rating on CV-6 (17.5 Turns)	2.65	2.64	.01	*
V_{24} - Mean of Altitude Deviations on CV-6 (17.5 Turns)	63.24	94.0	2.65	.23
V_{25} - Mean of AOA Control Ratings on CV-6 (17.5 Turns)	2.91	2.86	.15	*
V_{26} - Mean of Overall Repetition Ratings on CV-6 (17.5 Turns)	2.84	2.75	.28	*
V_{27} - CV-6 Gradeslip Ratings +	2.5	2.58	.20	**
V_{28} - 1st Repetition Overall Rating on CV-6 (21.5 Turns)	2.85	2.59	1.79	.14
V_{29} - Mean of Altitude Deviations on CV-6 (21.5 Turns)	59.2	78.1	1.03	.02

Table C-3. *(Continued)*

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,38)	Power Estimate
V_{30} - Mean of AOA Control Ratings on CV-6 (21.5 Turns)	2.97	2.8	1.47	.10
V_{31} - Mean of Overall Repetition Ratings on CV-6 (21.5 Turns)	2.95	2.77	1.30	.08

\+ Denotes variables derived from student mission gradeslip data; all other variables derived from BAM student performance questionnaire data.

* Denotes variables for which between group variances too small to compute power estimate.

** Denotes variables for which power estimate not computed due to better control group performances.

*** $p < .05$.

Table C-4. **Variables, Group Means and Statistical Power Estimates (alpha = .05) for Aircraft 30° DRR Data**

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,38)	Power Estimate
V_1 - Highest Repetition Rating obtained in A/C	3.25	3.05	1.03	.03
V_2 - Number of Repetitions to Proficiency	1.55	2.25	5.07***	.50
V_3 - 1st Repetition Overall Rating on CV-4	1.5	1.25	1.27	.08
V_4 - Mean of Airspeed Deviations on CV-4 Repetitions	9.94	13.1	.89	*
V_5 - Mean of Dive Angle Deviations on CV-4 Repetitions	2.55	3.47	2.36	.21
V_6 - Mean of Overall Repetition Ratings on CV-4 Repetitions	2.19	1.96	2.01	.16
V_7 - CV-4 Gradeslips Ratings +	1.9	1.8	.29	*
V_8 - 1st Repetition Overall Rating on CV-6	2.2	2.13	.08	*
V_9 - Mean of Airspeed Deviations on CV-6 Repetitions	3.62	3.44	.01	**
V_{10} - Mean of Dive Angle Deviations on CV-6 Repetitions	1.83	1.39	.60	**
V_{11} - Mean of Overall Repetition Ratings on CV-6 Repetitions	2.62	2.59	.03	*
V_{12} - CV-6 Gradeslip Ratings +	2.35	2.35	.00	*

\+ Denotes variables derived from student mission gradeslip data; all other variables derived from BAM student performance questionnaire data.
* Denotes variables for which between group variances too small to compute power estimate.
** Denotes variables for which power estimate not computed due to better control group performances.
*** $p < .05$.

Table C-5. **Variables, Group Means and Statistical Power Estimates (alpha = .05) for Aircraft 20° DRR Data**

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,38)	Power Estimate
V_1 - Highest Repetition Rating obtained in A/C	3.4	3.15	1.55	.11
V_2 - Number of Repetitions to Proficiency	1.35	1.35	.00	*
V_3 - 1st Repetition Overall Rating on CV-5	2.05	1.84	.84	*
V_4 - Mean of Airspeed Deviations on CV-5 Repetitions	5.5	8.29	2.69	.24
V_5 - Mean of Dive Angle Deviations on CV-5 Repetitions	1.71	1.49	.45	**
V_6 - Mean of Overall Repetition Ratings on CV-5 Repetitions	2.59	2.29	4.29***	.44
V_7 - CV-5 Gradeslips Ratings +	2.3	2.11	1.21	.07
V_8 - 1st Repetition Overall Rating on CV-6	2.4	2.24	.46	*
V_9 - Mean of Airspeed Deviation on DV-6 Repetitions	4.93	4.91	.01	**
V_{10} - Mean of Dive Angle Deviations on CV-6 Repetitions	1.45	1.04	.63	**
V_{11} - Mean of Overall Repetition Ratings on CV-6 Repetitions	2.77	2.61	.94	*
V_{12} - CV-6 Gradeslip Ratings +	2.45	2.26	1.53	.11

\+ Denotes variables derived from student mission gradeslip data; all other variables derived from BAM student performance questionnaire data.

* Denotes variables for which between group variances too small to compute power estimate.

** Denotes variables for which power estimate not computed due to better control group performances.

*** $p < .05$.

Table C-6. **Variables, Group Means and Statistical Power Estimates (alpha = .05) for Aircraft Target Tracking Data**

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,38)	Power Estimate
V_1 - Highest Repetition Rating obtained in A/C	2.7	2.53	.67	*
V_2 - Number of Repetitions to Proficiency	1.1	1.32	1.49	.10
V_3 - 1st Repetition Overall Rating	2.3	2.0	2.41	.21
V_4 - Mean of Overall Repetition Ratings	2.48	2.18	3.13	.30
V_5 - Gradeslip Ratings +	2.21	1.95	3.31	.32

\+ Denotes variables derived from student mission gradeslip data; all other variables derived from BAM student performance questionnaire data.

* Denotes variables for which between group variances too small to compute power estimate.

APPENDIX D: PHASE II TASK VARIABLES, GROUP MEANS, AND AIRCRAFT DATA F VALUES WITH STATISTICAL POWER ESTIMATES

Table D-1. **Variables, Group Means and Statistical Power Estimates (alpha = .05) for Aircraft Sortie BAM-1 LLN Data**

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df = 1,31)	Power Estimate
V_1 - Mean Number of Times Lost Over LLN Route Legs	.34	.39	.38	*
V_2 - Mean Time Deviation Over Route Checkpoints	16.8	17.5	.02	*
V_3 - Time Deviation on Last Checkpoint	20.4	16.9	.18	**
V_4 - Mean Altitude Deviation Over LLN Route Legs	328.1	173.1	3.27	**
V_5 - Mean Airspeed Deviation Over LLN Route Legs	16.9	9.7	10.25***	**
V_6 - Mean Course Control Rating Over LLN Route Legs	2.65	2.64	.00	*
V_7 - Gradesheet Ratings +	2.43	2.28	.34	**

+ Denotes variables derived from student mission gradeslip data; all other variables derived from LLN student performance questionnaire data.

* Denotes variables for which between group variances too small to compute power estimate.

** Denotes variables for which power estimates not computed due to better control group performances.

*** $p < .05$

Table D-2. **Variables, Group Means and Statistical Power Estimates (alpha = .05) for Aircraft Sortie SA-1 LLN Data**

Variable	$Mean_{exp}$	$Mean_{cont}$	F Value (df 1,31)	Power Estimate
V_1 - Mean Number of Times Lost Over LLN Route Legs	.23	.16	.44	**
V_2 - Mean Time Deviation Over Route Checkpoints	11.23	8.8	.94	**
V_3 - Time Deviation on Last Checkpoint	10.82	8.4	.65	**
V_4 - Mean Altitude Deviation Over LLN Route Legs	276.4	208.7	1.71	**
V_5 - Mean Airspeed Deviation Over LLN Route Legs	12.7	8.7	1.78	**
V_6 - Mean Course Control Rating Over LLN Route Legs	2.56	2.6	.22	**
V_7 - Gradesheet Ratings +	2.08	2.4	2.32	**

\+ Denotes variables derived from student mission gradeslip data; all other variables derived from LLN student performance questionnaire data.

* Denotes variables for which between group variances too small to compute power estimate.

** Denotes variables for which power estimates not computed due to better control group performances.

ME

-8